Journal of Pharmaceutical Analysis 12 (2022) 627–636

Contents lists available at ScienceDirect

Journal of Pharmaceutical Analysis

journal homepage: www.elsevier.com/locate/jpa

Original article

Multi-factor combined biomarker screening strategy to rapidly diagnose Alzheimer's disease and evaluate drug effect based on a rat model

Yanmeng Liu, Xinyue Zhang, Weiwei Lin, Nurmuhammat Kehriman, Wen Kuang, Xiaomei Ling*

The State Key Laboratory of Natural and Biomimetic Drugs and Department of Pharmaceutical Analysis, School of Pharmaceutical Sciences, Peking University, Beijing, 100191, China

ARTICLE INFO

Article history:
Received 19 August 2021
Received in revised form
11 April 2022
Accepted 13 April 2022
Available online 19 April 2022

Keywords:
Biomarker discovery
Metabolomics
Multi-factor combined biomarker screening strategy
Disease diagnosis
Drug evaluation

ABSTRACT

Alzheimer's disease (AD) represents the main form of dementia; however, valid diagnosis and treatment measures are lacking. The discovery of valuable biomarkers through omics technologies can help solve this problem. For this reason, metabolomic analysis using ultra-performance liquid chromatography coupled with quadrupole time-of-flight tandem mass spectrometry (UPLC-Q-TOF-MS) was carried out on plasma, hippocampus, and cortex samples of an AD rat model. Based on the metabolomic data, we report a multi-factor combined biomarker screening strategy to rapidly and accurately identify potential biomarkers. Compared with the usual procedure, our strategy can identify fewer biomarkers with higher diagnostic specificity and sensitivity. In addition to diagnosis, the potential biomarkers identified using our strategy were also beneficial for drug evaluation. Multi-factor combined biomarker screening strategy was used to identify differential metabolites from a rat model of amyloid beta peptide 1–40 ($A\beta_{1-40}$) plus ibotenic acid-induced AD (compared with the controls) for the first time; lysophosphatidylcholine (LysoPC) and intermediates of sphingolipid metabolism were screened as potential biomarkers. Subsequently, the effects of donepezil and pine nut were successfully reflected by regulating the levels of the abovementioned biomarkers and metabolic profile distribution in partial least squares-discriminant analysis (PLS-DA). This novel biomarker screening strategy can be used to analyze other metabolomic data to simultaneously enable disease diagnosis and drug evaluation.

1. Introduction

Alzheimer's disease (AD) represents the main form of dementia, and it poses a significant healthcare challenge. Currently, there are 50 million people with cognitive impairment worldwide and this number will increase to 152 million people by 2050 [1]. However, up to date, the diagnosis and treatment of AD are not sufficient [2]. Biomarkers can be used to diagnose diseases and evaluate the efficacy of treatments or drugs in target populations [3,4]. Therefore, the identification of valuable biomarkers is crucial for the diagnosis and treatment of AD.

Metabolomics, an effective platform for biomarker discovery, involves qualitative and quantitative metabolite measurements within biological systems [5,6]. Liquid chromatography combined with mass spectrometry (LC-MS) can produce massive amounts of metabolite information, which makes it a powerful technique for performing metabolomic analyses [7]. However, the vastness of LC-MS data, along with the diverse structures of metabolites, makes biomarker discovery difficult because it requires extensive data processing steps [8,9]. Biomarker screening is a key link in the data processing steps, and thus, several biomarker screening strategies have been developed in the field of metabolomics [10,11]. For example, Yun et al. [10] proposed a rank aggregation strategy to generate a "super" list that could reflect nine-variable importance analysis methods, thereby solving the problem of inconsistency among various variable ranking methods. However, these strategies focus only on the diagnostic capability of the screened biomarkers. Biomarkers can be classified into several types, such as predictive, diagnostic, pharmacodynamic, prognostic, safety, risk, and monitoring biomarkers. Pharmacodynamic biomarkers are assessed

Peer review under responsibility of Xi'an Jiaotong University.

* Corresponding author.
E-mail address: lingxm@bjmu.edu.cn (X. Ling).

https://doi.org/10.1016/j.jpha.2022.04.003

before and after drug dosing, thereby reflecting the effects of the drug on the disease [3]. Therefore, biomarkers that can simultaneously reflect multiple aspects of the disease are desirable.

Recently, a multi-factor combined biomarker screening strategy proposed by our group was able to identify the most discriminating metabolites for diagnosis due to its scientific selection of single-factor threshold and multi-factor combination [11]. Herein, we further explored whether biomarkers identified using this strategy can simultaneously diagnose AD and evaluate drug efficacy based on metabolomic data. In this study, a rat model of AD was established by the intrahippocampal injection of amyloid beta peptide 1–40 ($A\beta_{1-40}$) plus ibotenic acid (proven to be more representative of AD than Aβ or ibotenic acid alone) [12,13]. Then, a multi-factor combined biomarker screening strategy was used for biomarker discovery based on metabolomic data from the plasma, hippocampus, and cortex. As expected, 10, 7, and 6 potential biomarkers with high specificity and sensitivity for diagnosis were successfully identified in the plasma, hippocampus, and cortex samples, respectively. Subsequently, donepezil and a natural medicine, pine nut, were chosen for investigating the potential of the biomarkers identified using our strategy to evaluate drug efficacy. Donepezil is a drug that is approved by the U.S. Food and Drug Administration for the clinical treatment of AD. Pine nut has been reported to be helpful in AD treatment. Pine nut oil and its main component, pinolenic acid, have been shown to have antioxidant and anti-inflammatory effects [14,15], while oxidative stress and chronic inflammation are the main factors that accelerate the AD process. In addition, an active component of pine nut (UPNO-1) has been shown to ameliorate neurotoxicity and oxidative stress in an animal model of D-galactose-induced AD [16]. Taking advantage of the potential biomarkers identified using our strategy, the effects of drugs such as donepezil and pine nut were reflected by the regulation of overall metabolic profiles and metabolic biomarker levels. A multi-factor combined biomarker screening strategy could rapidly and accurately identify potential biomarkers beneficial for disease diagnosis as well as drug efficacy evaluation.

2. Materials and methods

2.1. Materials

$A\beta_{1-40}$ (sequence: DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGGVV) was obtained from AnaSpec (Fremont, CA, USA). Ibotenic acid (≥98%) was purchased from Cayman Chemical (Ann Arbor, MI, USA). Donepezil hydrochloride was purchased from National Medical Products Administration (Beijing, China). Korean pine nut was acquired from the Chensongshan Special Product Company (Yichun, China). Acetylcholine (ACh), acetylcholinesterase (AChE), and choline acetyltransferase (ChAT) were purchased from Andy Gene (Beijing, China). Acetonitrile and methanol (MS grade) were obtained from Thermo Fisher Scientific (Rockford, IL, USA). Formic acid (MS grade) was purchased from Sigma-Aldrich (St. Louis, MO, USA). All other chemicals and solvents used were of analytical grade.

2.2. Animals

Male Sprague-Dawley rats (200 ± 10 g) were obtained from the Laboratory Animal Institute of Peking University Health Science Center (Certificate No.: SYXK 2011–0039). The feeding conditions for the rats were a 12-h day and night cycle at room temperature (25 ± 2 °C) and humidity of 50% ± 5%. The rats were provided free access to food and water. The rats were randomly assigned to four groups (n = 13 per group), i.e., the control, AD model, donepezil, and pine nut groups. The $A\beta_{1-40}$ plus ibotenic acid AD rat model was established in accordance with a previously published procedure [12]. Briefly, the rats were anesthetized and fixed on a stereotaxic apparatus. The hippocampi of rats in the AD model, donepezil, and pine nut groups were bilaterally injected with 2 μL of mixed $A\beta_{1-40}$ ibotenic acid solution (coordinates: bregma 3.0, anterolateral 2.0, and dorsoventral 3 mm); the control group received 2 μL of saline. Then, the rats in the control, AD model, donepezil, and pine nut groups were intragastrically administered an equal volume of saline, saline, donepezil (1.5 g/kg), and pine nut suspension (2.25 g/kg), respectively, once per day for 1 month. The dose of donepezil used to treat rats in this study was based on the clinically recommended dosage for humans [12]. The dose of pine nut for treating rats was calculated based on the human dose advised by the "Compendium of Materia Medica". All animal experiments were approved by the Peking University Health Science Center Administration Committee.

2.3. Sample collection

The rats were first anesthetized with pentobarbital sodium, and their orbital venous blood was collected in heparin-containing Eppendorf tubes. Plasma samples were obtained by centrifuging the blood at 1000 r/min for 8 min at 4 °C. The rats were then sacrificed by cervical dislocation, and their hippocampi and cortices were quickly collected on ice. The three samples were then cryopreserved at −80 °C.

2.4. Histopathological and cholinergic system evaluation

The morphology of the hippocampal neurons was observed after paraffin-embedded section slicing and hematoxylin and eosin (H&E) staining. In brief, the rats (n = 3 per group) were anesthetized, and myocardial perfusion was performed using water and paraformaldehyde. Next, the brains were collected and stored in a fixative for 1 day. The fixed brains were embedded in paraffin and sectioned into 3–5 μm slices using a histotome (Leica, Germany). After H&E staining was performed, the hippocampus area was observed under light microscopy using a IX-71 light microscope (Olympus, Tokyo, Japan).

Enzyme-linked immunosorbent assay (ELISA) was performed to measure hippocampal ACh, AChE, and ChAT levels. The hippocampus was homogenized with phosphate-buffered saline (1:9, m/V) and then centrifuged at 12,000 r/min for 15 min at 4 °C. The supernatants were collected, and the hippocampal ACh, AChE, and ChAT levels were measured according to the manufacturer's instructions.

2.5. Sample preparation and LC-MS analysis

Prior to analysis, the plasma, hippocampus, and cortex samples were thawed on ice. Acetonitrile (150 μL) was added to the 50 μL of plasma sample; after 1 min of vortexing, the mixture was centrifuged at 12,000 r/min for 15 min at 4 °C twice, and then the supernatant was collected. The hippocampus and cortex samples were mixed with acetonitrile at a ratio of 1:5, vortexed for 1 min, and centrifuged at 12,000 r/min for 20 min at 4 °C twice; the supernatants were collected. Quality control (QC) samples were prepared by blending equivalent volumes of each sample.

Untargeted metabolomic analysis was performed in a positive ion mode on a Waters Xevo G2-Q/TOF mass spectrometer equipped with an ACQUITY UHPLC system (Waters Corporation, Milford, MA, USA). System repeatability and stability were tested through continuous injection of six QC replicates at the very beginning and implanting one QC sample between every ten real samples. Plasma,

hippocampus, and cortex samples (10 μL) were injected into a BEH C_{18} column (2.1 mm × 50 mm, 1.7 μm; Waters Corporation, Milford, MA, USA). The metabolites were measured at a flow rate of 0.4 mL/min according to the following LC gradient: 0–1 min, 85%–60% phase A (water containing 0.1% formic acid); 1–4 min, 60%–10% phase A; 4–5 min, 10% phase A. Phase B consisted of acetonitrile containing 0.1% formic acid. MS data were collected in the full scan mode with *m/z* scan range 50–1200 Da; leucine-enkephalin was utilized for real-time quality correction; desolvent gas flow rate was 600 L/h; the ion source temperature was 110 °C; and the capillary and extraction cone voltages were 3.5 and 2.0 kV, respectively.

2.6. Data and statistical analysis

The raw metabolomic data were processed using MarkerLynx4.1 XS (Waters Corporation, Milford, MA, USA) for peak detection and alignment. The parameters were set as follows: retention time range, 0–6 min; mass range, 50–1200 Da; retention time window, 0.1 min; mass window, 0.05 Da; mass tolerance, 0.02 Da; intensity threshold, 100; peak width at 5% height, 1.0 s; peak-to-peak baseline noise, 0.0; and noise elimination level, 6. Multivariate statistical analyses, including principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA), were performed using SIMCA 13.0 demo software (Umetrics Corporation, Malmo, Sweden). The normal distribution of data was checked using the Kolmogorov-Smirnov test and Shapiro-Wilk test, while the homogeneity of variance of data was checked using Levene's test. Student's *t*-test and receiver operating characteristic (ROC) curve analysis were performed using SPSS software (IBM Corporation, Armonk, NY, USA). Potential biomarkers were identified based on the match of precise mass and MS^2 information to databases (such as Human Metabolome Database (HMDB) and MassBank). The metabolic pathways were analyzed using the kyoto encyclopedia of genes and genomes database resource (https://www.kegg.jp/).

2.7. Multi-factor combined biomarker screening strategy

The strategy consisted of three steps, i.e., method verification, single-factor screening, and multi-factor combined screening. The method verification step involved the selection of the initial variables from the primitive variables. Relative standard deviation (RSD) value of the variable's peak intensity in QC samples <20% was used as the selection criterion. The number of initial variables was labeled as N_0.

The single-factor screening step was performed to obtain the optimal threshold for the four factors, namely, the variable importance in projection (VIP), fold change (FC), area under the ROC curve (AUROC), and -ln(*P* value). Their thresholds are labeled as *a*, *b*, *c*, and *d*, respectively, while N_a, N_b, N_c, and N_d represent the numbers of variables satisfying *a*, *b*, *c*, and *d*, respectively. The minimum thresholds were based on the general cutoff of the factors, which were 1.0, 1.0, 0.70, and 3.0, for *a, b, c,* and *d*, respectively. Regression analyses were then performed to explore the relationship between N_a/N_0 and *a*, N_b/N_0 and *b*, N_c/N_0 and *c*, and N_d/N_0 and *d*. The maximum thresholds were acquired from the regression functions, and they were the corresponding values when $N/N_0 = 0.01$. In this study, five thresholds were set up for each factor, and they were equally spaced. Next, the variables satisfying the threshold of each factor were selected, and corresponding PLS-DA analysis was performed. The R^2X, R^2Y, and Q^2 values represent the PLS-DA model's explanation, classification, and forecasting capabilities, respectively. Overall, the sum of R^2X, R^2Y, and Q^2 values (SUM) was chosen as the index for optimal threshold selection. The best threshold for each factor was obtained based on the maximum SUM.

The multi-factor combined screening step was used to further screen the variables by comparing the multi-factor combinations. For the four factors, there were 11 combinations (combs 1–11). Eleven multi-factor combinations and corresponding variable sets were obtained according to the optimal thresholds of the four factors, and then PLS-DA analysis was performed. Consistent with the single-factor screening step, the combination with maximal SUM was chosen as the best combination. The variables in the best combination were the potential biomarkers.

3. Results and discussion

3.1. Pathomorphological changes

Pyramidal cell necrosis in the hippocampal CA1 region is an important pathological basis for learning and memory dysfunction in AD [17]. H&E staining is the most commonly used method of pathological tissue section staining, in which abnormal pyramidal neurons tend to shrink with pyknotic and hyperchromatic nuclei [18]. As shown in Fig. 1A, neurons in the CA1 region of the control group showed a normal morphology with clear boundaries and lightly stained nucleoli. Marked morphological changes, including karyopyknosis and dark staining, were visualized in the AD model group, indicating the occurrence of neuronal lesions. Karyopyknosis was ameliorated after donepezil and pine nut treatment; nevertheless, the neurons still displayed a looser arrangement compared to the neurons of the control group. In summary, AD modeling by $A\beta_{1-40}$ plus ibotenic acid administration resulted in severe neuronal damage, and both donepezil and pine nut exhibited a certain degree of neuroprotection. Pathomorphological results provided a preliminary basis for selecting donepezil and pine nut as drugs to be included in the efficacy evaluation.

3.2. Damage to the cholinergic system

ACh is closely involved in learning and memory processes, and the loss of cholinergic neurons in the brain is one of the most important pathological features of AD [19,20]. ChAT and AChE are the two key enzymes that generate ACh and degrade ACh, respectively. ChAT, ACh, and AChE form a relatively balanced and mutually constrained cycle under physiological conditions. Relative to the Ach levels in the control group, the ACh levels in the AD model group were significantly lower (Fig. 1B; $P < 0.05$). Donepezil treatment significantly increased the ACh concentration (Fig. 1B; $P < 0.05$), whereas pine nuts did not significantly change the ACh levels. For AChE, $A\beta_{1-40}$ plus ibotenic acid injection significantly increased AChE activity, while both donepezil and pine nut treatment inhibited this increase significantly (Fig. 1C; $P < 0.05$). The ELISA results of AChE confirmed the cholinesterase inhibitor nature of donepezil and showed that pine nut had a similar effect as donepezil. Additionally, ChAT activity was significantly lower after AD modeling but significantly elevated after donepezil treatment (Fig. 1D; $P < 0.05$). In conclusion, impaired Ach, AChE, and ChAT in the AD model group were all ameliorated by donepezil treatment, reflecting its good efficacy in AD. AChE activity elevation due to AD modeling was inhibited by pine nut treatment, indicating that pine nut could serve as a new therapeutic medicine through the AChE inhibition effect, similar to the widely used clinical drugs, such as donepezil, galantamine, and rivastigmine [21,22]. The results of cholinergic system further suggested that donepezil and pine nut could be used in the efficacy evaluation.

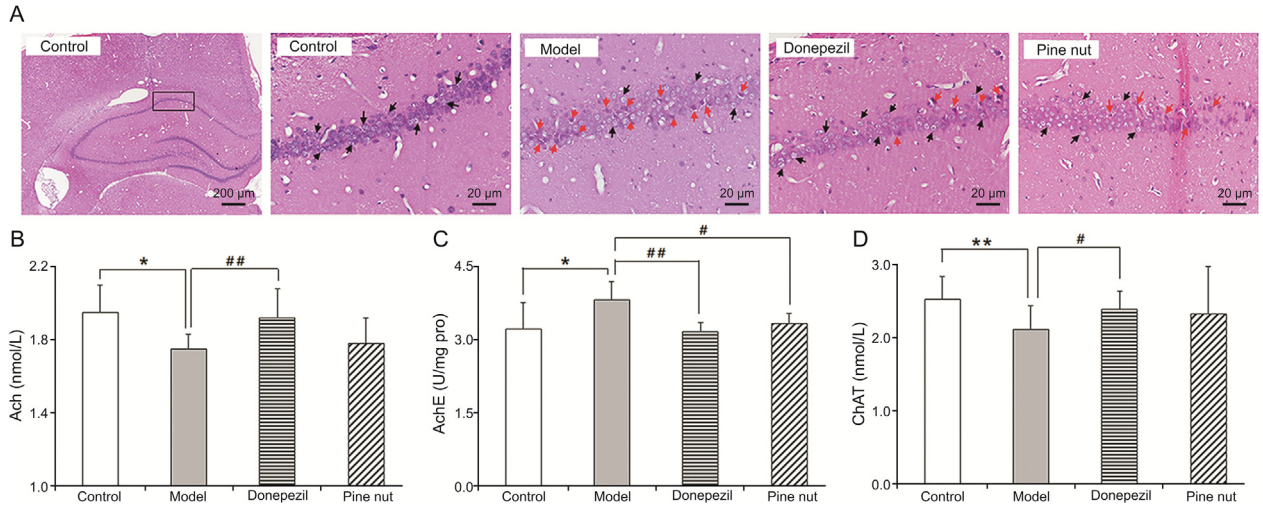

Fig. 1. (A) Hematoxylin and eosin (H&E) staining images of the pyramidal cells in the hippocampus CA1 region of the control, AD model, donepezil, and pine nut groups. The black arrowhead represents normal cells and the red arrowhead represents distorted cells. (B) Acetylcholine (Ach) level, (C) acetylcholinesterase (AchE) activity, and (D) choline acetyltransferase (ChAT) activity of control, AD model, donepezil, and pine nut groups (mean ± standard deviation, $^{*}P < 0.05$ presented comparing with model group, $^{\#}P < 0.05$ and $^{\#\#}P < 0.01$ presented comparing with drug (donepezil or pine nut) group).

3.3. Biomarker discovery using multi-factor combined biomarker screening strategy

Non-targeted metabolomics was used to identify differential metabolites, and the total ion chromatograms (TIC) are shown in Fig. S1. In addition, all six QC samples clustered together in the PCA score plots, demonstrating the satisfactory reliability and repeatability of the metabolomic data for the plasma, hippocampus, and cortex samples (Fig. S2). Next, a multi-factor combined biomarker screening strategy was employed to identify differential metabolites in a rat model of $A\beta_{40}$ plus ibotenic acid-induced AD (compared with the controls), which has been achieved for the first time, based on metabolomic data.

First, 332, 185, and 267 initial variables were obtained for the plasma, hippocampus, and cortex samples, respectively, using the method verification step. Second, the initial variables were screened by comparing the different thresholds of VIP, FC, AUROC, and −ln(*P* value). Regression analyses were carried out to explore the relationship between N_a/N_0 and a, N_b/N_0 and b, N_c/N_0 and c, *and* N_d/N_0 and d in the three samples (Figs. 2 and S3). The R^2 values of the regression functions were all greater than 0.9, reflecting high-fitting reliability. The minimum thresholds were 1.0, 1.0, 0.70, and 3.0, for *a, b, c,* and *d*, respectively; the maximum thresholds were acquired from the regression functions, and they were the corresponding values when $N/N_0 = 0.01$. Five thresholds were set up for each factor in the present study. Table 1 shows the corresponding values of a and N_a, b and N_b, c and N_c, d and N_d in the plasma, hippocampus, and cortex samples. Next, PLS-DA analysis was performed to determine the optimal threshold for each factor. The score plots of PLS-DA models performed by the variable sets selected through different thresholds of the four factors are shown in Fig. S4 (plasma), Fig. S5 (hippocampus), and Fig. S6 (cortex). For the plasma sample, the R^2X, R^2Y, and Q^2 values at different thresholds of the four factors are shown in Figs. 2E–H. As for VIP, with increasing a, the R^2X and Q^2 values were first downregulated and then upregulated, while the R^2Y value was downregulated (Fig. 2E). The change rules of FC, AUROC, and −ln(*P* value) were similar; R^2X increased but R^2Y and Q^2 decreased as the threshold increased (Figs. 2F–H). It can be seen that the biggest R^2X, R^2Y, and Q^2 values were obtained at different thresholds. Overall, the SUM was chosen as the index for optimal threshold selection. The SUM values at different thresholds of the four factors are listed in Table 1. With the maximum SUM, 2.1, 0.87, and 5.7 were chosen as the optimal thresholds of *b*, *c*, and *d*, respectively. There was one exception to the VIP. The largest SUM was obtained when $a = 3.8$; however, there was an outlier that was outside the 95% confidence interval in the PLS-DA score plot when $a = 3.8$ (Fig. S4E). For the remaining four thresholds, larger SUM values were obtained at $a = 1.0$ and $a = 1.7$, while the control and AD model specimens were separated better at 1.7 than at 1.0 (Figs. S4A and B). Consequently, we selected 1.7 as the optimal value of *a* for the plasma samples. In summary, the optimal values of *a*, *b*, *c*, and *d* for the plasma samples were 1.7, 2.1, 0.87, and 5.7, respectively. Similarly, we acquired the optimal thresholds for the hippocampus and cortex samples according to the largest SUM (Fig. S7 and Table 1). As a result, 1.0, 1.0, 0.70, 3.8 were the optimal *a*, *b*, *c*, *d* for the hippocampus samples; 1.0, 1.0, 0.85, 6.2 were the optimal *a*, *b*, *c*, *d* for the cortex samples.

Finally, the second-step-obtained variable sets were further screened by comparing multi-factor combinations. The numbers of variables satisfying the requirements of multi-factors in the plasma, hippocampus, and cortex samples are shown in the Wayne plots (Figs. 3A–C). Table 2 shows eleven multi-factor combinations and their corresponding numbers of variables in the plasma, hippocampus, and cortex samples. For plasma samples, set5 and set10, and set8 and set11 were the same. Consistent with the single-factor screening step, the combination with maximal SUM was chosen as the best combination. The results revealed that when the PLS-DA model was performed using set6, the SUM was the highest (Table 2). Therefore, comb6 was the best multi-factor combination, and the variables in set6 were the expected potential biomarkers. Similarly, we acquired the best combinations of the hippocampus and cortex samples according to the largest SUM (Table 2). As a result, set5 (or set6, set10) was the best combination for the hippocampus samples; set6 (or set10) was the best combination for cortex samples. In summary, the best combination containing 10, 7, and 6 variables was finally acquired for plasma, hippocampus, and cortex samples, respectively (Figs. 3D–F). The PLS-DA score plots and R^2X, R^2Y, and Q^2 values of the best combination in the three samples are shown in Figs. 3D–F, while the PLS-DA score plots and parameters of the remaining ten combinations are displayed in Fig. S8 (plasma),

Table 1
The thresholds of four factors, the numbers of corresponding variables, and the sum of R^2X, R^2Y, and Q^2 values (SUM) of the partial least squares-discriminant analysis (PLS-DA) models performed by the corresponding variables in plasma, hippocampus, and cortex samples.

Sample	N_0	VIP			FC			AUROC			−ln(*P* value)		
		a	N_a	SUM	*b*	N_b	SUM	*c*	N_c	SUM	*d*	N_d	SUM
Plasma	332	1.0	68	1.93	1.0	282	1.94	0.70	157	2.07	3.0	98	2.18
		1.7	29	1.87	2.1	17	2.12	0.76	95	2.19	3.9	55	2.23
		2.4	17	1.74	3.2	7	2.09	0.81	54	2.29	4.8	29	2.27
		3.1	8	1.54	4.3	4	1.97	0.87	20	2.40	5.7	13	2.36
		3.8	4	2.35	5.5	3	1.95	0.92	3	2.20	6.7	4	2.23
Hippocampus	185	1.0	23	2.46	1.0	132	2.00	0.70	25	2.51	3.0	15	2.44
		1.9	9	1.67	1.8	16	1.80	0.75	15	2.45	3.8	7	2.65
		2.7	7	1.61	2.6	8	1.89	0.80	9	1.91	4.6	4	2.32
		3.6	4	1.63	3.4	8	1.89	0.85	3	1.96	5.4	2	2.31
		4.4	3	1.84	4.3	2	1.85	0.89	2	2.22	6.1	2	2.31
Cortex	267	1.0	46	2.37	1.0	218	2.00	0.70	88	2.15	3.0	68	2.11
		1.7	14	2.14	2.0	15	1.59	0.75	51	2.52	4.1	37	2.56
		2.4	8	2.15	3.0	9	1.54	0.80	38	2.30	5.1	23	2.43
		3.1	5	2.11	4.0	4	1.41	0.85	19	2.70	6.2	7	2.46
		3.8	3	2.14	5.0	3	1.39	0.91	4	2.60	7.2	4	2.49

N_0: the number of initial variables; VIP: variable importance in projection; FC: fold change; AUROC: area under receiver operator characteristic curve; *a*, *b*, *c*, and *d*: the threshold of VIP, FC, AUROC and −ln(*P* value), respectively; N_a, N_b, N_c, and N_d: the numbers of corresponding variables that met the requirements of the thresholds; *P* value: AD model group compared with control group.

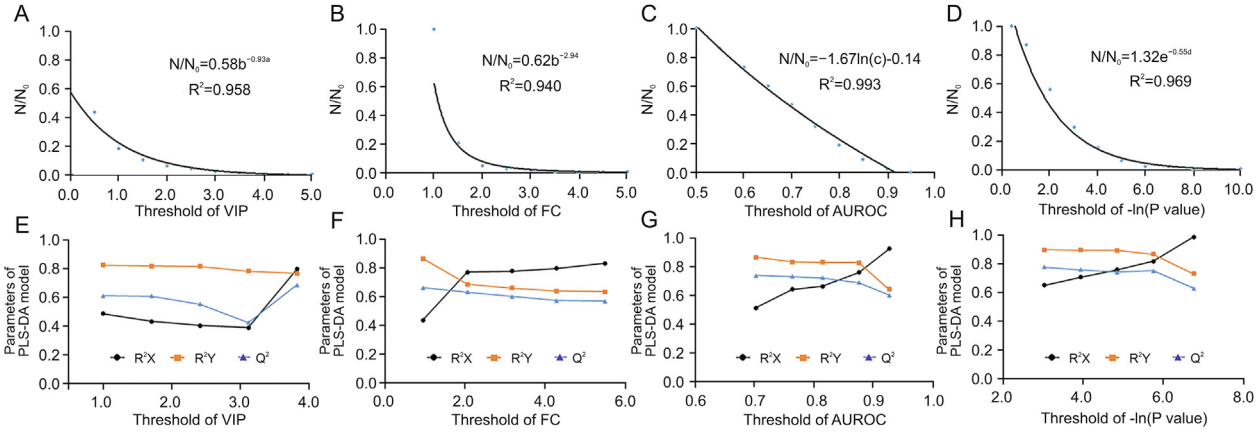

Fig. 2. (A–D) Regression curves between N/N_0 and the factor's threshold in the plasma sample. *N* is the number of variables that meets the factor's threshold. N_0 is the number of initial variables. *a*, *b*, *c*, and *d* represent the thresholds of variable importance in projection (VIP), fold change (FC), area under receiver operator characteristic curve (AUROC), and −ln(*P* value), respectively. (E–H) The R^2X, R^2Y, and Q^2 values of the partial least squares-discriminant analysis (PLS-DA) models performed by the variable sets selected through considering the thresholds of (E) VIP, (F) FC, (G) AUROC , and (H) −ln(*P* value) in the plasma sample.

Fig. S9 (hippocampus), and Fig. S10 (cortex). As shown in Figs. 3D–F, the variables of the best combination could successfully distinguish AD model rats from control rats with excellent parameters in all three samples. The obtained R^2X, R^2Y, and Q^2 values were 0.83, 0.85, and 0.81 for plasma; 0.99, 0.81, and 0.71 for hippocampus; and 0.82, 0.94, and 0.90 for cortex. Moreover, the AUROC of the selected variables by our strategy was all greater than 0.85 (Figs. 3G–I), which demonstrated their high specificity and sensitivity for diagnosis. In summary, 10, 7, and 6 potential biomarkers with high diagnostic specificity and sensitivity were successfully obtained using our strategy for the plasma, hippocampus, and cortex samples, respectively.

VIP>1.0, FC > 1.2, and $P < 0.05$ were commonly used in the usual procedure of biomarker screening. Using the usual procedure, 34, 6, and 18 variables were obtained for the plasma, hippocampus, and cortex samples, respectively (Figs. 4A, D, and G). The R^2X, R^2Y, and Q^2 values of the PLS-DA model performed by these variables were as follows: 0.63, 0.85, and 0.70, respectively, for the plasma; 1.00, 0.45, and 0.33, respectively, for the hippocampus; and 0.87, 0.70, and 0.54, respectively, for the cortex. In contrast, 10, 7, and 6 variables were obtained using our strategy for plasma, hippocampus, and cortex samples, respectively (Figs. 3D–F). The variables selected by our strategy can provide an excellent distinction between the control and AD model groups with better parameters. The obtained R^2X, R^2Y, and Q^2 values by our strategy were 0.83, 0.85, and 0.81, respectively, for plasma; 0.99, 0.81, and 0.71, respectively, for hippocampus; and 0.82, 0.94, and 0.90, respectively, for cortex. In particular, the overlapped control and model groups in the hippocampus and cortex samples were successfully separated using our strategy. In summary, compared with the usual procedure, our strategy greatly reduced the number of screened biomarkers and significantly improved the explanation, classification, and forecasting capabilities of the PLS-DA model. There were two possible reasons for these observations. The first was the broad screening criteria for the usual procedure leading to large numbers of biomarkers being identified, with significant time and effort required to identify them and elucidate their metabolic pathways. The other was that the AUROC was not taken into consideration in

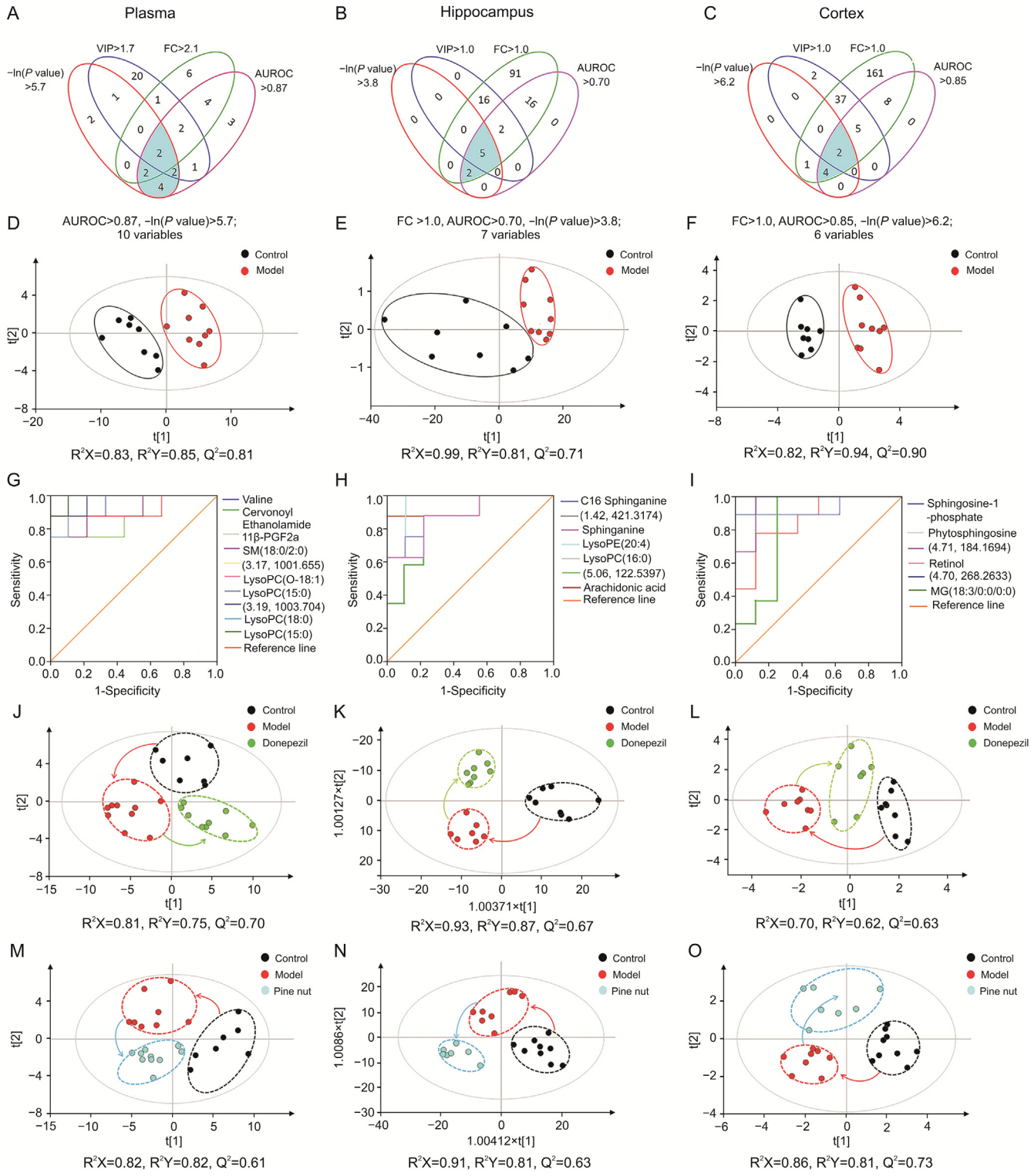

Fig. 3. (A–C) Wayne plots of the numbers of variables meeting the requirements of multi-factor in plasma, hippocampus, and cortex samples. The best multi-factor combination is marked in blue in the Wayne plot. (D–F) The score plots and parameters of PLS-DA model performed by the best combination in plasma, hippocampus, and cortex samples. Black and red circles represent the control and AD model groups, respectively. (G–I) The receiver operator characteristic curves of the variables of the best combinations in the plasma, hippocampus, and cortex samples. (J–L) The score plots of PLS-DA models performed by the selected variables of control, AD model, and donepezil groups in the plasma, hippocampus, and cortex samples. Therapeutic effect of donepezil was reflected by the curative direction towards the control group. Black, red, and green circles represent the control, AD model and donepezil groups, respectively. The red arrowhead represents the modeling direction, and the green arrowhead represents the curative direction. (M–O) The score plots of PLS-DA models performed by the selected variables of control, AD model, and pine nut groups. The therapeutic effect of pine nut was reflected by the curative direction towards the control group. Black, red, and light blue circles represented control, AD model, and pine nut groups, respectively. The red arrowhead represents the modeling direction, and the light blue arrowhead represents the curative direction.

Table 2
The numbers (*N*) of variables and the SUM of the eleven multi-factor combinations in plasma, hippocampus, and cortex samples.

No.	Multi-factor combination	Plasma		Hippocampus		Cortex	
		N	SUM	*N*	SUM	*N*	SUM
1	VIP and FC	5	2.18	23	1.81	44	2.40
2	VIP and AUROC	7	2.34	7	1.83	7	2.41
3	VIP and −ln(*P* value)	5	2.36	5	1.92	2	2.47
4	FC and AUROC	10	2.17	25	2.15	19	2.60
5	FC and −ln(*P* value)	4	2.21	7	2.51	7	2.46
6	AUROC and −ln(*P* value)	10	2.49	7	2.51	6	2.66
7	VIP, FC, and AUROC	4	2.17	7	1.83	7	2.41
8	VIP, FC, and −ln(*P* value)	2	1.98	5	1.92	2	2.47
9	VIP, AUROC, and −ln(*P* value)	4	2.46	5	1.92	2	2.47
10	FC, AUROC, and −ln(*P* value)	4	2.21	7	2.51	6	2.66
11	VIP, FC, AUROC, and −ln(*P* value)	2	1.98	5	1.92	2	2.47

VIP: variable importance in projection; FC: fold change; AUROC: area under the receiver operator characteristic curve.

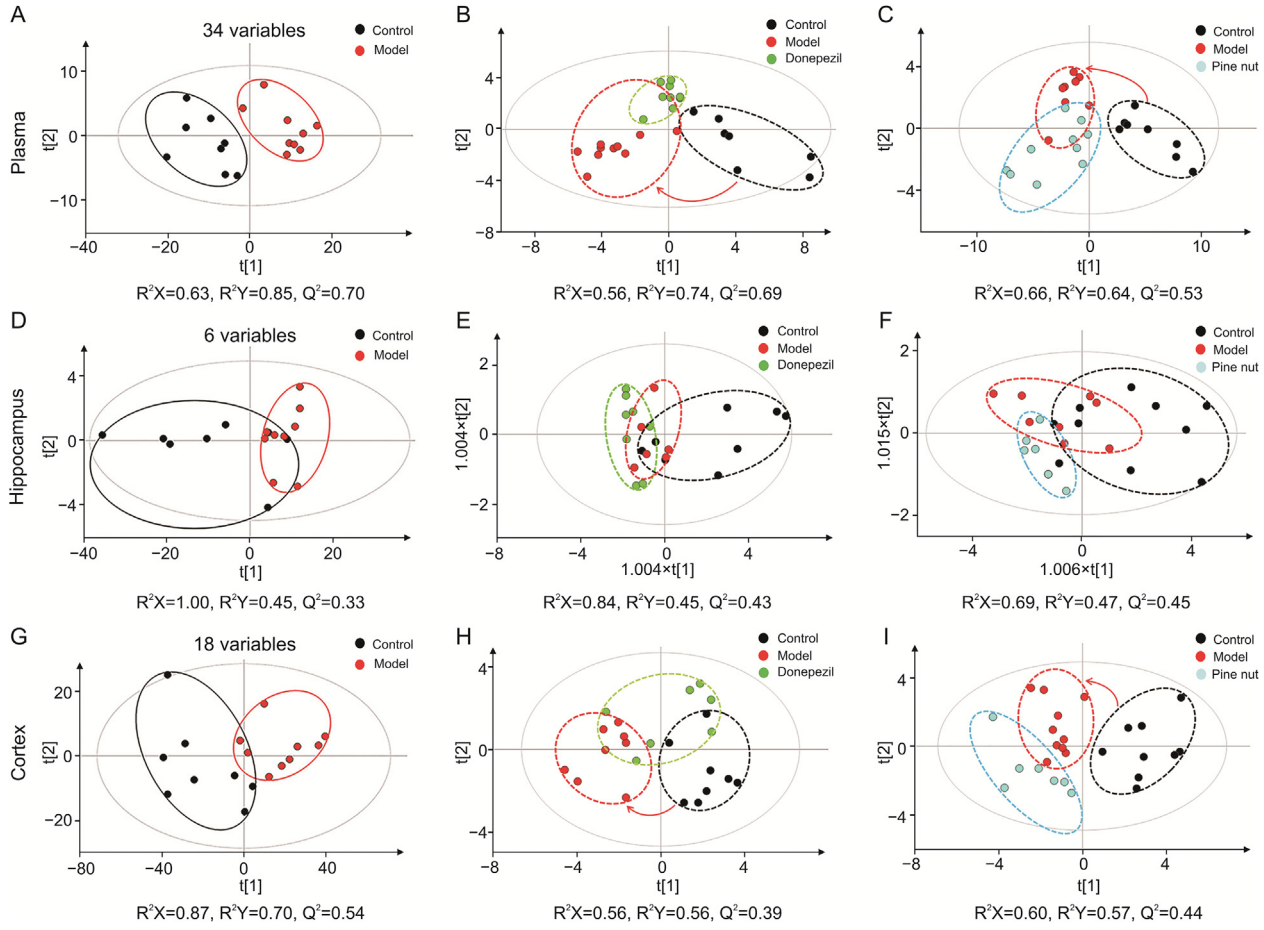

Fig. 4. (A, D, and G) The score plots and parameters of the PLS-DA models performed by the variables selected by usual procedure in plasma, hippocampus, and cortex samples. Black and red circles represent the control and AD model groups, respectively. (B, E, and H) The score plots of the PLS-DA models performed by the selected variables of the control, AD model, and donepezil groups in plasma, hippocampus, and cortex samples. Black, red, and green circles represent the control, AD model, and donepezil groups, respectively. The red arrowhead represents the modeling direction. (C, F, and I) The score plots of PLS-DA models performed by the selected variables of control, AD model, and pine nut groups in plasma, hippocampus, and cortex samples. Black, red, and light blue circles represent the control, AD model, and pine nut group, respectively. The red arrowhead represents the modeling direction.

the usual procedure, while the AUROC of a variable is an important test of its grouping ability.

3.4. Biomarkers discovered by multi-factor combined biomarker screening strategy for drug evaluation

Using the abovementioned variables selected by our strategy, grouping of the control, AD, and drug groups in three samples was obtained by PLS-DA analysis (Fig. 3). Compared with the AD model group, both the donepezil and pine nut groups were distributed more closely to the control group in all three samples, reflecting their benign effects on AD. Moreover, the R^2X, R^2Y, and Q^2 values of the PLS-DA models were all greater than 0.6, demonstrating that the built models possessed powerful explanatory, grouping, and predictive abilities. In contrast, exploiting the variables screened by the usual procedure, the control, AD model, and drug groups could

Table 3
The names, metabolic pathways, and changing trends of the biomarkers in the plasma, hippocampus, and cortex samples.

Sample	Retention time (min)	Measured *m/z* (Da)	Calculated *m/z* (Da)	Error of *m/z* (Da)	Identity	Metabolic pathway	FC1*	FC2	FC3
Plasma	0.3	118.0865	118.0868	0.0003	Valine	–	1.78	–	0.75*
	1.9	373.2733	373.2742	0.0009	Cervonoyl ethanolamide	–	0.41	3.45*	–
	1.9	355.2629	355.2584	0.0045	11β-PGF2a	Arachidonic acid metabolism	0.40	3.24*	–
	2.9	506.3556	506.3558	0.0002	SM (18:0/2:0)	Sphingolipid metabolism	0.51	–	1.47*
	3.3	508.3754	508.3766	0.0012	LysoPC(O-18:1)	Glycerophospholipid metabolism	0.59	1.31*	1.48*
	3.6	482.3237	482.3246	0.0009	LysoPC(15:0)	Glycerophospholipid metabolism	0.69	1.27*	–
	3.6	524.3712	524.3716	0.0004	LysoPC(18:0)	Glycerophospholipid metabolism	0.86	–	–
	3.7	550.3866	550.3816	0.0050	LysoPC(20:2)	Glycerophospholipid metabolism	0.72	1.47*	1.50*
Hippocampus	1.3	274.2740	274.2746	0.0006	C16 sphinganine	Sphingolipid metabolism	0.82	–	–
	1.6	302.3053	302.3059	0.0006	Sphinganine	Sphingolipid metabolism	0.78	–	–
	1.8	502.2925	502.2933	0.0008	LysoPC(20:4)	Glycerophospholipid metabolism	1.17	–	0.86*
	2.1	496.3397	496.3403	0.0006	LysoPC(16:0)	Glycerophospholipid metabolism	1.19	–	–
	3.4	305.2473	305.2480	0.0007	Arachidonic acid	Arachidonic acid metabolism	1.23	0.73*	0.78*
Cortex	0.9	380.2213	380.2265	0.0052	Sphingosine-1-phosphate	Sphingolipid metabolism	0.21	7.64*	5.62*
	1.3	318.2999	318.3008	0.0009	Phytosphingosine	Sphingolipid metabolism	0.67	–	–
	3.4	287.2369	287.2375	0.0006	Retinol	–	1.43	–	–
	3.5	353.2669	353.2692	0.0023	MG (18:3/0:0/0:0)	–	2.15	0.49*	0.44*

FC1: AD model group compared with control group; FC2: donepezil group compared with AD model group; FC3: pine nut group compared with AD model group. *$P < 0.05$.

not be separated in the PLS-DA score plots (Fig. 4). In addition, the curative direction of donepezil and pine nut groups towards the control group could not be reflected by the usual procedure.

8, 5, and 4 variables were identified among the 10, 7, and 6 selected variables of the plasma, hippocampus, and cortex samples, respectively. The names, metabolic pathways, and changing trends of the identified variables are shown in Table 3. It is shown that a total of three metabolic pathways were disturbed after AD modeling. Sphingolipid metabolism was perturbed in all three samples, while glycerophospholipid and arachidonic acid metabolisms were disturbed in two samples (plasma and hippocampus). It was also shown that the metabolic pathways of the three samples were crossed and the plasma biomarkers could cover the biomarkers of hippocampus and cortex samples, which indicated that plasma samples could replace hippocampal and cortical samples to elucidate the disturbed metabolic pathway of AD using our strategy. This holds some significance for human beings, which means that it is possible to replace the unobtainable hippocampus and cortex with accessible plasma for AD diagnosis. In addition, there were 2, 2, and 2 variables unidentified in the plasma, hippocampus, and cortex samples, respectively. Their retention times, observed *m/z* values, and changing trends are listed in Table S1.

As shown in Table 3, five metabolites involved in sphingolipid metabolism were impaired after AD modeling. Donepezil rescued the decreased sphingosine-1-phosphate level, while pine nut reversed the downregulation of SM (18:0/2:0) and sphingosine-1-phosphate. Glycerophospholipid metabolism is mainly focused on lysophosphatidylcholine (LysoPC). In comparison with the control group, four plasma LysoPCs were downregulated in the AD model group. Donepezil reversed three of them, i.e., LysoPC(O-18:1), LysoPC(18:0), and LysoPC(20:2); and pine nut rescued two, i.e., LysoPC(O-18:1) and LysoPC(20:2). LysoPC(16:0) and LysoPC(20:4) levels in the hippocampus were elevated after modeling, and the increased LysoPC(20:4) level was rescued by pine nut treatment. Moreover, as for arachidonic acid metabolism, the upregulated hippocampal arachidonic acid level induced by modeling was reversed by both donepezil and pine nut. Metabolites that were regulated by both donepezil and pine nut treatment included sphingosine-1-phosphate, LysoPC(O-18:1), LysoPC(20:2), and arachidonic acid.

In this study, after AD modeling, sphingolipid metabolism was perturbed in the plasma, hippocampus, and cortex samples. Sphingolipid metabolism disorders might contribute to AD neuropathological changes by regulating amyloid precursor protein (APP) processing and disrupting synaptic activity [23]. Compared with the control group, sphinganine was decreased in the hippocampus of AD model rats in this study. It is well known that sphinganine can be converted into ceramide, while an increase in ceramide can promote the production of IL-2 and IL-6, thus causing neuroinflammation [23]. Sphingosine-1-phosphate is a neuroprotective factor that has been reported to decline with increasing Braak stage in the brain region [24]. Consistently, the level of sphingosine-1-phosphate was reduced in the cortex of AD rats compared to the controls. Glycerophospholipid metabolism is mainly focused on LysoPC. LysoPC can sustain the natural glycerophospholipid composition of the neuron membrane. Downregulation of LysoPC may damage brain membrane phospholipids, and lower levels of LysoPC have been reported in serum samples of AD patients versus control subjects [25]. Consistently, we discovered four downregulated LysoPCs in the plasma sample of the AD model group using our strategy. Arachidonic acid metabolism is strongly correlated with inflammation and oxidative stress, two important mechanisms related to AD [26,27]. Arachidonic acid can regulate leukocyte chemotaxis and inflammatory cytokine production, which can lead to an inflammatory reaction [26]. An increase in arachidonic acid was observed in the brain of an AD transgenic mouse model [28]. Consistent with this, elevation of arachidonic acid levels was also found in the hippocampus of AD model rats. In summary, the three metabolic pathways enriched by the screened biomarkers were the common pathways affected by AD. This means that compared with the usual procedure, our strategy identified the key metabolic pathways in AD with fewer biomarkers and significantly increased the efficiency of biomarker screening, thus accelerating the subsequent time-consuming biomarker identification process.

Drugs such as donepezil and pine nut were evaluated based on the regulation of the overall metabolic profiles and screened metabolic biomarker levels. The therapeutic effects of donepezil and pine nut were reflected in the curative direction towards the control group in the PLS-DA score plots and by regulating screened biomarker levels closer to the control group. Sphingosine-1-phosphate in the sphingolipid metabolism pathway is a neuroprotective agent that decreases gradually as AD progresses [24]. Herein, both donepezil and pine nut treatment rescued the decreased sphingosine-1-phosphate level, which reflected their

good regulatory effect on sphingolipid metabolism. Moreover, donepezil and pine nut adjusted three and two impaired LysoPCs, respectively, in the plasma sample. This indicated that donepezil and pine nut can promote the stability of cell membranes to different degrees. The adjusted action of donepezil and pine nut on arachidonic acid also reflects their regulatory effects on neuroinflammation and oxidative stress aspects of AD pathological processes. Donepezil and pine nut have similar regulatory effects on some biomarkers, such as sphingosine-1-phosphate, two LysoPCs, and arachidonic acid. AChE ELISA results showed that pine nut administration inhibited AChE significantly, similar to the AChE inhibitor donepezil. The similar biomarker regulation effects of donepezil and pine nuts were consistent with their similar AChE inhibition effect, further illuminating the validity of using biomarkers obtained by the strategy to evaluate drug effects.

Nevertheless, not all impaired biomarkers were rescued by donepezil and pine nut, and the biomarkers regulated by donepezil and pine nut were not always the same. For example, in the plasma sample, donepezil and pine nut did not normalize the down-regulated LysoPC(18:0), whereas LysoPC(15:0) was only regulated by donepezil, and SM(18:0/2:0) was only adjusted by pine nut. Due to the rigorous biomarker screening criteria of our strategy, the screened biomarkers were the most discriminating metabolites between the control and AD model groups. Donepezil did not actually cure AD in the clinic, although it is a widely used drug in the management of AD [29]. In this study, donepezil did not restore all potential biomarkers screened out by our strategy to normal levels consistently.

4. Conclusions

In the present study, we proposed a multi-factor combined biomarker screening strategy to rapidly and accurately discover potential biomarkers that are beneficial for both disease diagnosis and drug evaluation. Our strategy was applied to discover differential metabolites from $A\beta_{1-40}$ plus ibotenic acid-induced AD model compared with the controls for the first time based on plasma, hippocampus, and cortex metabolomic data. As a result, 10, 7, and 6 potential biomarkers with high diagnostic specificity and sensitivity were identified for the plasma, hippocampus, and cortex samples, respectively. The obtained potential biomarkers were mainly LysoPCs and intermediates of sphingolipid metabolism, including LysoPC(O-18:1), LysoPC(15:0), LysoPC(18:0), LysoPC(20:2), LysoPC(16:0), sphinganine, sphingosine-1-phosphate, and phytosphingosine. Compared with the usual procedure, our strategy greatly reduced the number of screened biomarkers and significantly improved the explanatory, classification, and forecasting capabilities of the PLS-DA model. Drug efficacy evaluation was performed using the screened biomarkers. The effects of donepezil and pine nut were successfully reflected by the regulation of biomarker levels and metabolic profile distribution of PLS-DA. Therefore, the biomarkers identified using our strategy are diagnostic and pharmacodynamic biomarkers. We hope that a multi-factor combined biomarker screening strategy can be used to analyze clinical metabolomic data to quickly and accurately identify biomarkers and evaluate drug efficacy.

CRediT author statement

Yanmeng Liu: Data curation, Writing - Original draft preparation; **Xinyue Zhang:** Investigation; **Weiwei Lin:** Supervision; **Nurmuhammat Kehriman:** Validation; **Wen Kuang:** Data curation; **Xiaomei Ling:** Conceptualization.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No.: 81673392).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpha.2022.04.003.

References

[1] Alzherimer's Disease International, World Alzheimer Report 2018. https://www.alz.co.uk/research/world-report-2018. (Accessed 24 December 2021).
[2] J. Ding, K.L. Davis-Plourde, S. Sedaghat, et al., Antihypertensive medications and risk for incident dementia and Alzheimer's disease: A meta-analysis of individual participant data from prospective cohort studies, Lancet Neurol. 19 (2020) 61–70.
[3] A.K. Arakaki, J. Skolnick, J.F. McDonald, Marker metabolites can be therapeutic targets as well, Nature 456 (2008) 443.
[4] Y.N. Wang, Y. Sun, C. Lau, et al., Duplex microRNAs assay based on target-triggered universal reporter hybridization, J. Pharm. Anal. 8 (2018) 265–270.
[5] S. Murugesu, Z. Ibrahim, Q.U. Ahmed, et al., Identification of α-glucosidase inhibitors from Clinacanthus nutans leaf extract using liquid chromatography-mass spectrometry-based metabolomics and protein-ligand interaction with molecular docking, J. Pharm. Anal. 9 (2019) 91–99.
[6] Y.-Y. Chen, J. Shen, Y.-P. Tang, et al., Elucidating the interaction of Kansui and licorice by comparative plasma/tissue metabolomics and a heatmap with relative fold change, J. Pharm. Anal. 9 (2019) 312–323.
[7] T. Liu, R. Li, Y. Cui, et al., Metabonomic analysis of plasma biochemical changes in pyrexia rats after treatment with Gegenqinlian decoction, aspirin and itraconazole by UHPLC-FT-ICR-MS, J. Pharm. Anal. 10 (2020) 581–587.
[8] E. Gorrochategui, J. Jaumot, S. Lacorte, et al., Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: overview and workflow, Trac. Trends Anal. Chem. 82 (2016) 425–442.
[9] L. Yi, N. Dong, Y. Yun, et al., Chemometric methods in data processing of mass spectrometry-based metabolomics: a review, Anal. Chim. Acta 914 (2016) 17–34.
[10] Y.-H. Yun, B.-C. Deng, D.-S. Cao, et al., Variable importance analysis based on rank aggregation with applications in metabolomics for biomarker discovery, Anal. Chim. Acta 911 (2016) 27–34.
[11] C. Li, J. Zhang, R. Wu, et al., A novel strategy for rapidly and accurately screening biomarkers based on ultraperformance liquid chromatography-mass spectrometry metabolomics data, Anal. Chim. Acta 1063 (2019) 47–56.
[12] W. Lin, J. Zhang, Y. Liu, et al., Studies on diagnostic biomarkers and therapeutic mechanism of Alzheimer's disease through metabolomics and hippocampal proteomics, Eur. J. Pharmaceut. Sci. 105 (2017) 119–126.
[13] Y. Li, H.Q. Qin, Q.S. Chen, et al., Behavioral and neurochemical effects of the intrahippocampal co-injection of β-amyloid protein1-40 and ibotenic acid in rats, Int. J. Neurosci. 114 (2004) 1521–1531.
[14] X. Chen, Y. Zhang, Z. Wang, et al., In vivo antioxidant activity of *Pinus koraiensis* nut oil obtained by optimised supercritical carbon dioxide extraction, Nat. Prod. Res. 25 (2011) 1807–1816.
[15] S.-J. Chen, L.-T. Chuang, J.-S. Liao, et al., Phospholipid incorporation of non-methylene-interrupted fatty acids (NMIFA) in murine microglial BV-2 cells reduces pro-inflammatory mediator production, Inflammation 38 (2015) 2133–2145.
[16] J. Zhang, W. Lin, R. Wu, et al., Mechanisms of the active components from Korean pine nut preventing and treating D-galactose-induced aging rats, Biomed. Pharmacother. 103 (2018) 680–690.
[17] D.R. Sharma, W.Y. Wani, A. Sunkaria, et al., Quercetin attenuates neuronal death against aluminum-induced neurodegeneration in the rat hippocampus, Neuroscience 324 (2016) 163–176.
[18] Q. Quan, J. Wang, X. Li, et al., Ginsenoside Rg1 decreases Aβ(1-42) level by upregulating PPARγ and IDE expression in the hippocampus of a rat model of Alzheimer's disease, PLoS One 8 (2013), e59155.
[19] A. Easton, V. Douchamps, M. Eacott, et al., A specific role for septohippocampal acetylcholine in memory? Neuropsychologia 50 (2012) 3156–3168.
[20] B.M. Bennett, J.N. Reynolds, G.T. Prusky, et al., Cognitive deficits in rats after forebrain cholinergic depletion are reversed by a novel NO mimetic nitrate ester, Neuropsychopharmacology 32 (2007) 505–513.
[21] C.G. Parsons, W. Danysz, A. Dekundy, et al., Memantine and cholinesterase inhibitors: complementary mechanisms in the treatment of Alzheimer's disease, Neurotox. Res. 24 (2013) 358–369.
[22] D.S. Woodruff-Pak, M.J. Tobia, X. Jiao, et al., Preclinical investigation of the

functional effects of memantine and memantine combined with galantamine or donepezil, Neuropsychopharmacology 32 (2007) 1284–1294.
[23] N.J. Haughey, V.V.R. Bandaru, M. Bae, et al., Roles for dysfunctional sphingolipid metabolism in Alzheimer's disease neuropathogenesis, Biochim. Biophys. Acta 1801 (2010) 878–886.
[24] T.A. Couttas, N. Kain, B. Daniels, et al., Loss of the neuroprotective factor sphingosine 1-phosphate early in Alzheimer's disease pathogenesis, Acta Neuropathol. Commun. 2 (2014), 9.
[25] Y. Cui, X. Liu, M. Wang, et al., Lysophosphatidylcholine and amide as metabolites for detecting Alzheimer disease using ultrahigh-performance liquid chromatography-quadrupole time-of-flight mass spectrometry-based metabonomics, J. Neuropathol. Exp. Neurol. 73 (2014) 954–963.
[26] W.-Q. Zhang, Y.-L. Hua, M. Zhang, et al., Metabonomic analysis of the anti-inflammatory effects of volatile oils of *Angelica sinensis* on rat model of acute inflammation, Biomed. Chromatogr. 29 (2015) 902–910.
[27] M.A. Balboa, J. Balsinde, Oxidative stress and arachidonic acid mobilization, Biochim. Biophys. Acta Mol. Cell Biol. Lipids 1761 (2006) 385–391.
[28] R.O. Sanchez-Mejia, J.W. Newman, S. Toh, et al., Phospholipase A2 reduction ameliorates cognitive deficits in a mouse model of Alzheimer's disease, Nat. Neurosci. 11 (2008) 1311–1318.
[29] A.C. Baakman, R. Alvarez-Jimenez, G. Loewen, et al., No synergistic effect of subtherapeutic doses of donepezil and EVP-6124 in healthy elderly subjects in a scopolamine challenge model, Alzheimers. Dement. (N Y) 5 (2019) 89–98.